

FIG.1A

GGCACGAGGCTTCTGGCCAGGGAACGTGGAAGCGCACCGACAGGGATCCGGCCAGGAG 60

GGCGAGTGAAAGAAGGAATCAGAAAGGAAGGAGTTAACAAATAATAAAACAGCCTG 120

AGCCACGGCTGGAGAGACCGAGACCCGGCGCAAGAGAGCGCAGCCTTAGTAGGAGAGGAA 180

CGCGAGACCGGGCAGCGCAGAGCGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 240

AGCGCGCGCAGCTGAGACCGGGCGCGCAGCGCCAGCCCTCAGGGGGGCTCACAAGTCAG 300

CGCCCAAGCAAGTCAAGCGACAGCGCTCGTCTTCGCCCGAACTGATGCGCTGCAAACGCC 360

M R C K R R 6

GGCTCAACTTCAGCGGCTTTGGCTACAGCCTGCCGACGACGAGCGCCGGCCGTGGCGC 420

L N F S G F G Y S L P Q Q Q P A A V A R 26

FIG.1B

GCCGCAACGAGCGCGGCAACCGGCTCAAGTTGGTCAACCTGGGCTTTGCCACCCTTC 480
R N E R E R N R V K L V N L G F A T L R 46
GGAGCACGTCCCAACGGCGGCCACAAGAATGAGTAAGGTGAGACACTGCGCT 540
E H V P N G A A N K K M S K V E T L R S 66
CGGCGGTGAGTACATCCGGCGGCTGCAGCAGCTGCTGGACGAGCATGACGCGGTGAGCG 600
A V E Y I R A L Q Q L L D E H D A V S A 86
CCGCCCTTCCAGGCAAGCGTCTGTGCCCCACCATCTCCCCAACTACTCCAACGACTTGA 660
A F Q A G V L S P T I S P N Y S N D L N 106
ACTCCATGGCCGGCTCGCCGGTCTCATCTACTCGTCGGACGAGGGCTCTTACGACCCGC 720
S M A G S P V S S Y S S D E G S Y D P L 126
TCAGCCCCGAGGAGCAGGAGCTTCTCGACTTCACCAACTGGTCTGAGGGGCTCGGCCTG 780
S P E E Q E L L D F T N W F * 140
GTCAGGCCCTGGTGCGAATGACTTTGGAAGCAGGGTGATCGCACCAACCTGCATCTTTAG 840
TGCTTTCTTGTCAGTGGCGTTGGGAGGGGAGAAAGGAAAGAAAGAAAGAAAGAGAAGA 900

FIG.1C

AGAAGAAAAGAGAAGAAAACGAAACAGTCAACCAACCCATCGCCAACTAAGC 960
GAGGCATGCCCTGAGAGACATGGCTTTCAGAAAACGGGAAGCGCTCAGAACAGTATCTTG 1020
CACTCCAATCATTCACGGAGATATGAAGCAACTGGGACCTGAGTCAATGCCAAAATG 1080
CAGCTTGTGTGCAAAAGCAGTGGGCTCCTGGCAGAAGGAGCAGCACCGGTTATAGTA 1140
ACTCCCATCACCTCTAACACGCACAGCTGAAAGTCTTGCTGGGTCCCTTCACCTCCCC 1200
GCCCTTCTTAGAGTGCAGTCTTAGCCCTTAGAAAACGAGTTGGTGCTTTCGTCTCAG 1260
TAGCCCCCACCACAATAAGCTGTAGACATTGGTTTACAGTGAAACTATGCTATTCTCAGC 1320
CCTTTGAAACTCTGCTTCTCCTCCAGGGCCCGATTCCCAACCCCATGGCTTCCCTCACA 1380

FIG.1D

CTGTCYTTTCTACCATYTTTCATTATAGAATGCTTCCAATCTTTTGGAATTTTTTATTAT 1440
AAAAAATCTATTGTATCTATCCCTAACCAAGTTCGGGATATATTAGATATTTTGTACA 1500
TAAGAGAGAAAGAGAGAGAAAAATTTATAGAAGTTTGTACAATGGTTTAAATGTGTA 1560
TATCTTGATACTTTAACATGTAATGCTATTACCTCTGCATATTTAGATGTAGTTCAC 1620
CTTACAAC TGCAATTTCCCTATGTGGTTTGTAAAGAACTCTCCCTATAGGTGAGATCA 1680
AGAGGCCACCAGTTGTACTTCAGCACCAATGTGTCTTACTTTATAGAATA GTTGTAAATG 1740
TATTAA TGATGTTATTAAATACTGTTCAGAAGAACAAGTTTATGCAGCTACTGTCCAA 1800
ACTCAAAGTGGCAGCCAGTTGGTTTGTATAGGTTGCCCTTTTGGAGATTTCTATTACTGCC 1860

FIG.1E

TTTTTTTTCTTACTGTTTATTACAACCTTACAATAATGTATAACCCCTGTTTATACA 1920
AACTAGTTTCGTAATAAACCTTTTCCCTTTTAAATG 1960

FIG.2A

GGCACGAGGCTTCTGCCAGGGAACGTGGAAGGCCACCACAGGATCCGGCCAGGAG 60
GGCGAGTGAAGAAGAAATCAGAAAGGAAGGAGTTAACAAATATAAACAACACCTG 120
AGCCACGGCTGGAGAGACCAGACCCGGCGCAAGAGAGCCGACCTTAGTAGAGAGGAA 180
CGCGAGACGGCGCAGAGCGCGTTCAGCACTTCTGCTGCTTCTGCTTTT 240
TCTTAGAAACAAGAGCGCGCAGCGCCTCACACGGAGCGCCACGCGAGGCTCCCG 300
AAGCCAACCCGCGAAGGGAAGGGAAGGGAAGGAGCGCGCTGCAGGGAAGAGAAA 360
AGCATTTTCACTTTTGTGCTCCCACTTAAGAAGTCTCCCGGGGATTTGTATATATT 420
TTTAACTTCCGTCAGGCGTCCCGCTTCATATTTCCCTTTCTTCCCTCTCTGTTCTGCA 480

FIG.2B

CCCAAGTGGTCAACCTGGGCTTTGCCACCCTTCGGGAGCACGTCCCAACGGCGGCC 540

AACAGAGAATGAGTAAGTGGAGACACTGCGCTCGGCGTCGAGTACATCCGCGCCTG 600

 M S K V E T L R S A V E Y I R A L 17

CAGCAGCTGCTGGACGAGCATGACCGCGGTGAGCGCGCCTTCCAGGACGGCGTCTGCG 660

Q Q L L D E H D A V S A A F Q A G V L S 37

CCCACCATTCTCCCCCAACTACTCCAACGACTTGAACCTCATGGCCGGCTCGCCGCTCTCA 720

P T I S P N Y S N D L N S M A G S P V S 57

TCCTACTCGTCGGACGAGGGCTCTTACGACCCGCTCAGCCCCGAGGAGCAGAGCTTCTC 780

S Y S S D E G S Y D P L S P E E Q E L L 77

GACTTCACCAACTGTTCTGAGGGGCTCGGCCCTGGTCAGGCCCTGGTGCGAATGACTTT 840

D E T N W F * 83

GGAAGCAGGGTGATCGCACACCTGCATCTTTAGTGCTTCTTGTCAGTGCGGTTGGAG 900

GGGGAGAAAAGGAAAAGAAAAGAAAAGAGAAGAGAAGAAAAGAGAAGAAAAGAAC 960

FIG.2C

GAAAACAGTCAACCAACCCCATCGCCAACTAAGCGAGGCATGCCCTGAGAGACATGGCTTT 1020
CAGAAACGGGAAGCGCTCAGAACAGTATCTTTGCACTCCAATCATTCACGGAGATATGA 1080
AGAGCAACTGGGACCTGAGTCAATCGCAAAATGCAGCTTGTGTGCAAAAAGCAGTGGGCT 1140
CCTGGCAGAAGGAGCAGCACACCGCTTATAGTAACTCCCATCACCCTTAACACGCACAG 1200
CTGAAGTTCTTGGCTCGGGTCCCTTCACCTCCCCGCCCTTCTTAGAGTGCAGTTCTTAG 1260
CCCTCTAGAAACGAGTTGGTGTCTTTCGTCTCAGTAGCCCCCACCCTCAATAAGCTGTAGA 1320
CATTGGTTTACAGTGAAACTATGCTATTCTCAGCCCTTTGAAACTCTGCTTCTCCTCCAG 1380
GGCCCCGATTCCCAAAACCCCATGGCTTCCCTCACACTGTCTTTCTACCATTTTCATTATA 1440

FIG.2D

GAATGCTCCAATCTTTGTGAATTTTATTATAAAAAATCTATTTGTATCTATCCTAA 1500
CCAGTTCGGGATATATTAGATATTTTGTACATAAGAGAGAGAGAGAAAAATTT 1560
ATAGAAGTTTGTACAATGTTTAAATGTGTATATCTTGATACTTTAACATGTAATGC 1620
TATTACCTCTGCATATTTAGATGTGAGTTCACCTTACAACCTGCAATTTCCCTATGTG 1680
GTTTGTAAAGAACTCTCCTCATAGGTGAGATCAAGAGGCCACCAAGTTGTACTTCAGCAC 1740
CAATGTCTTACTTTATAGAAATGTTGTTAATGTAATTAAATACTGTT 1800
CAAGAAGAACAAAGTTTATGCAGCTACTGTCCAACCTCAAAGTGCGACCAAGTTGGTTT 1860
GATAGGTTGCCCTTTGGAGATTTCTATTACTGCCCTTTTCTTACTGTTTATTACAA 1920

FIG.2E

ACTTACAATAATGTATAACCCCTGTTTATACAACTAGTTTCGTAATAAACTTTTC 1980

CTTTTAAATG 1994

FIG.3A

GGCACGAGGCTTCTGCCAGGGAACGTGGAAGCGCACCGACAGGGATCCGGCCAGGAG 60

GGCGAGTGAAAGAAGGAATCAGAAAGGAAGGAGTTAACAAATAATAAAACAGCCTG 120

AGCCACGGCTGGAGAGACCAGAACCCGGCGCAAGAGAGCGCAGCCTTAGTAGAGAGGAA 180

CGCGAGACGGGGCAGAGCGCGTTCAGCACTGACTTTTGCTGCTTCTGCTTTT'TT'T 240

TCTTAGAAACAAGAAGCGCCAGCGCAGCCTCACACGCGAGCGCCACGGAGGCTCCCG 300

AAGCCAACCCGGGAAGGAGGGGAGGAGGAGCGCGCTGCAAGGAGGAGGAAAA 360

AGCATTTTCACTTTT'TTGCTCCCACTT'AAGAACTTCCCGGGATTTGTATATA'TT 420

TTTAACTTCCGTCAGGGCTCCCGCTCAATATTTCCTTT'TCTTCCCTCTCTGTTCCTGCA 480

FIG.3B

CCCAAGTCTCTGTGTCCCTCGGGGGCCCGACCTCGCGTCCCGGATCGCTCTGA 540

TTCCGGACTCCTTGGCCCGCGCTGCGCATGGAAGCTCTGCCAAGATGGAGCGGCGG 600

M E S S A K M E S G G 11

CGCCGGCCAGCAGCCCCAGCCGCGAGCCCCAGCAGCCCTTCTGCCGCGCCGAGCCCTGTTT 660

A G Q Q P Q P Q Q P F L P P A A C F 31

CTTTGCCACGGCCGAGCCCGCGGGCCGCGAGCCCGCGAGCGGCGAGCGCGCA 720

F A T A A A A A A A A A Q S A Q 51

GCAGCAGCAGCAGCAGCAGCAGCAGCAGCGCGCGAGCTGAGACCGGGCGCGA 780

Q Q Q Q Q Q Q Q Q A P Q L R P A A D 71

CGGCCAGCCCTCAGGGGGGCTCACAAGTCAGCGCCCAAGCAAGTCAAGCGACGCGCTC 840

G Q P S G G G H K S A P K Q V K R Q R S 91

GTCTTCGCCCGAACTGATGGCTGCAAAAGCCGGCTCAACTTCAGCGGCTTGGCTACAG 900

S S P E L M R C K R R L N F S G F G Y S 111

CCTGCCGCGAGCAGCAGCTGCTGGACGAGCATGACGCGGTGAGCGCGCCTTCCAGGCAAG 960

L P Q Q Q L L D E H D A V S A A F Q A G 131

FIG.3C

CGTCCTGTGCCCAACCATCTCCCCCACTACTCCAACGACTTGAACCTCCATGGCCGGCTC 1020
V L S P T I S P N Y S N D L N S M A G S 151
GCCGCTCTCATCCTACTCGTCGGACGAGGGCTCTTACGACCCGCTCAGCCCGAGAGCA 1080
P V S S Y S S D E G S Y D P L S P E E Q 171
GGAGCTTCTCGACTTCACCAACTGGTTCTGAGGGGCTCGGCCTGGTCAGGCCCTGGTGCG 1140
E L L D E T N W F * 180
AATGACTTTGGAAGCAGGGTGATCGCACCAACCTGCATCTTTAGTGCTTTCTTGTCAGTG 1200
GCCTTGGGAGGGGAGAAAGGAAAGAAAGAAAGAAAGAGAAGAGAAGAAAGAGAAGA 1260
AGAAAAAAACGAACAACAGTCAACCAACCCCATCGCCAACTAAGCGAGGCATGCCCTGAGAG 1320
ACATGGCTTTCAGAAAAACGGGAAGCGCTCAGAACAGTATCTTGGACTCCAATCATTCAC 1380
GGAGATATGAAGACAACTGGGACCTGAGTCAATGGGCAAAATGCAGCTTGTGTGCAAAA 1440

FIG.3D

GCAGTGGGCTCCTGGCAGAAGGAGCAGCACGCCGTTAAGTAACCTCCCATCACCCTCTA 1500
ACACGCACAGCTGAAGAATTCTTGCTGGGTCCCTTCACCTCCCCGCCCTTCTTAGAGTG 1560
CAGTTCTTAGCCCTCTAGAAACGAGTTGGTGTCTTTCGTCTCAGTAGCCCCCACCCTCAAT 1620
AAGCTGTAGACATGGTTTACAGTGAACCTATGCTATTCTCAGCCCTTTGAAACTCTGCT 1680
TCTCCTCCAGGGCCGATTCCCAAAACCCATGGCTTCCCTCACACTGTCTTTTCTACCAT 1740
TTTCATTATAGAAATGCTTCCAATCTTTGTGAATTTTATTAAAAAATCTATTTGTA 1800
TCTATCCTAACCAAGTTCGGGATATATTAAGATATTTTGTACATAAGAGAAAGAGAG 1860
AGAAAAAATTATAGAGATTTTGTACAATAAGTTTAAATGTGTATATCTTGATACTTTAA 1920

FIG.3E

CATGTAATGCTATTACCTCTGCATATTTAGATGTGTAGTTCACCTTACAACCTGCAATT 1980
TCCCTATGTGTTTGTAAAGAACTCTCCTCATAGGTGAGATCAAGAGGCCACCAGTTGT 2040
ACTTCAGCACCAATGTCTTACTTTATAGAATGTTGTTAATGTATTATGATGTTAT 2100
AAATACTGTTCAAGAAGAACAAGTTTATGCAGCTACTGTCCAACCTCAAAGTGGCAGCC 2160
AGTTGGTTTGTATAGGTTGCCCTTTTGGAGATTCTATPACTGCCCTTTTCTTACTGT 2220
TTTATTACAACCTTACAATAATATGTATAACCCCTGTTTATACAACTAGTTTCGTAATA 2280
AAACTTTTTCCTTTTAAATAATG 2304

FIG.4A

1

60

SACHV1 GGCACGAGGCTTCTG GCCAGGGAACGTGGA AGGCGCACCACAGG GATCCGGCCAGGAG
SACHV2 GGCACGAGGCTTCTG GCCAGGGAACGTGGA AGGCGCACCACAGG GATCCGGCCAGGAG
SACHV3 GGCACGAGGCTTCTG GCCAGGGAACGTGGA AGGCGCACCACAGG GATCCGGCCAGGAG
SACH GGCACGAGGCTTCTG GCCAGGGAACGTGGA AGGCGCACCACAGG GATCCGGCCAGGAG

61

120

SACHV1 GGCAGTGAAGAAG GAAATCAGAAGGAA GCGAGTTAACAATA ATTAATAACAGCCTG
SACHV2 GGCAGTGAAGAAG GAAATCAGAAGGAA GCGAGTTAACAATA ATTAATAACAGCCTG
SACHV3 GGCAGTGAAGAAG GAAATCAGAAGGAA GCGAGTTAACAATA ATTAATAACAGCCTG
SACH GGCAGTGAAGAAG GAAATCAGAAGGAA GCGAGTTAACAATA ATTAATAACAGCCTG

FIG.4B

121

180

SACHV1 AGCCACGGCTGGAGA GACCGAGACCCGGCG CAAGAGAGCGCAGCC TTAGTAGGAGAGGAA
SACHV2 AGCCACGGCTGGAGA GACCGAGACCCGGCG CAAGAGAGCGCAGCC TTAGTAGGAGAGGAA
SACHV3 AGCCACGGCTGGAGA GACCGAGACCCGGCG CAAGAGAGCGCAGCC TTAGTAGGAGAGGAA
SACH AGCCACGGCTGGAGA GACCGAGACCCGGCG CAAGAGAGCGCAGCC TTAGTAGGAGAGGAA

181

240

SACHV1 CGCGAGACGGCGCAG CGC-----
SACHV2 CGCGAGACGGCGCAG AGCGCGTTCAGCACT GACTTTTGCTGCTGC TTCCTGCTTTT
SACHV3 CGCGAGACGGCGCAG AGCGCGTTCAGCACT GACTTTTGCTGCTGC TTCCTGCTTTT
SACH CGCGAGACGGCGCAG AGCGCGTTCAGCACT GACTTTTGCTGCTGC TTCCTGCTTTT

FIG.4C

241

300

SACHV1

SACHV2 TCTTAGAACAAGAA GCGCCAGCGGCAGC CTCACACGCGAGCGC CACGCGAGGCTCCCG

SACHV3 TCTTAGAACAAGAA GCGCCAGCGGCAGC CTCACACGCGAGCGC CACGCGAGGCTCCCG

SACH TCTTAGAACAAGAA GCGCCAGCGGCAGC CTCACACGCGAGCGC CACGCGAGGCTCCCG

301

360

SACHV1

SACHV2 AAGCCAACCCGCGAA GCGAGGAGGGGAGGG AGGAGGAGCGCGCT GCAGGAGGAGAAAA

SACHV3 AAGCCAACCCGCGAA GCGAGGAGGGGAGGG AGGAGGAGCGCGCT GCAGGAGGAGAAAA

SACH AAGCCAACCCGCGAA GCGAGGAGGGGAGGG AGGAGGAGCGCGCT GCAGGAGGAGAAAA

FIG.4D

361

420

SACHV1

AGCATTTTCACTTTT TTTGCTCCCACTCTA AGAAGTCTCCGGGG ATTTGTATATATT

SACHV3 AGCATTTTCACTTTT TTTGCTCCCACTCTA AGAAGTCTCCGGGG ATTTGTATATATT

SACH AGCATTTTCACTTTT TTTGCTCCCACTCTA AGAAGTCTCCGGGG ATTTGTATATATT

421

480

SACHV1

TTTAACTTCCGTCAG GGCTCCCGCTTCATA TTTCCTTTCTTTCC CTCCTGTTCCTGCA

SACHV3 TTTAACTTCCGTCAG GGCTCCCGCTTCATA TTTCCTTTCTTTCC CTCCTGTTCCTGCA

SACH TTTAACTTCCGTCAG GGCTCCCGCTTCATA TTTCCTTTCTTTCC CTCCTGTTCCTGCA

FIG.4E

481

540

SACHV1

SACHV2 CCCAAGTT

SACHV3 CCCAAGTTCCTCTG TGTCCCCCTCGCGG CCCCGACCTCGCGT CCCGGATCGCTCTGA

SACH CCCAAGTTCCTCTG TGTCCCCCTCGCGG CCCCGACCTCGCGT CCCGGATCGCTCTGA

541

600

SACHV1

SACHV2

SACHV3 TTCGGGACTCCTTG GCCGCCGCTGGGCAT GGAAGCTCTGCCAA GATGAGAGCGGCGG

SACH TTCGGGACTCCTTG GCCGCCGCTGGGCAT GGAAGCTCTGCCAA GATGAGAGCGGCGG

FIG.4F

601

660

SACHV1

SACHV2

SACHV3

SACH

661

720

SACHV1

SACHV2

SACHV3

SACH

AGAGCGCGCA

CGCCGGCCAGCAGCC CCAGCCGCAGCCCCA GCAGCCCTTCCTGCC GCCCGCAGCCTGTTT
CGCCGGCCAGCAGCC CCAGCCGCAGCCCCA GCAGCCCTTCCTGCC GCCCGCAGCCTGTTT
CTTGGCCACGGCCGC AGCCGGGGGGGGCCGC AGCCGGCCGACAGCGGC AGCGCAGAGCGCGCA
CTTGGCCACGGCCGC AGCCGGGGGGGGCCGC AGCCGGCCGACAGCGGC AGCGCAGAGCGCGCA

FIG.4G

721

780

SACHV1 GCAGCAGCAGCAGCA GCAGCAGCAGCAGCA GCAGGCCGCCAGCT GAGACCGCGGCCGA

SACHV2 -----

SACHV3 GCAGCAGCAGCAGCA GCAGCAGCAGCAGCA GCAGGCCGCCAGCT GAGACCGCGGCCGA

SACH GCAGCAGCAGCAGCA GCAGCAGCAGCAGCA GCAGGCCGCCAGCT GAGACCGCGGCCGA

781

840

SACHV1 CGGCCAGCCCTCAGG GGGCGGTCACAAGTC AGGCCCAAGCAAGT CAAGCGACAGCGCTC

SACHV2 -----

SACHV3 CGGCCAGCCCTCAGG GGGCGGTCACAAGTC AGGCCCAAGCAAGT CAAGCGACAGCGCTC

SACH CGGCCAGCCCTCAGG GGGCGGTCACAAGTC AGGCCCAAGCAAGT CAAGCGACAGCGCTC

FIG.4H

841

900

SACHV1 GTCTTCGCCCGAACT GATGCGCTGCAACG CCGGCTCAACTTCAG CGGCTTTGGCTACAG

SACHV2 -----

SACHV3 GTCTTCGCCCGAACT GATGCGCTGCAACG CCGGCTCAACTTCAG CGGCTTTGGCTACAG

SACH GTCTTCGCCCGAACT GATGCGCTGCAACG CCGGCTCAACTTCAG CGGCTTTGGCTACAG

901

960

SACHV1 CCTGCCGCAAGCA GCCGGCCGCCGTGC GCGCGCAACGAGCG CGAGCGCAACCGCGT

SACHV2 -----

SACHV3 CCTGCCGCAAGCA GC-----

SACH CCTGCCGCAAGCA GCCGGCCGCCGTGC GCGCGCAACGAGCG CGAGCGCAACCGCGT

FIG.4I

961

1020

SACHV1 CAAGTTGGTCAACCT GGGCTTTGCCACCCT TCGGAGCAGCTCCC CAACGGCGCGGCCAA

SACHV2 -----GGTCAACCT GGGCTTTGCCACCCT TCGGAGCAGCTCCC CAACGGCGCGGCCAA

SACHV3 -----

SACH CAAGTTGGTCAACCT GGGCTTTGCCACCCT TCGGAGCAGCTCCC CAACGGCGCGGCCAA

1021

1080

SACHV1 CAAGAAGATGAGTAA GGTGAGACACTGCG CTCGGCGTCAAGTA CATCCGCGCGCTGCA

SACHV2 CAAGAAGATGAGTAA GGTGAGACACTGCG CTCGGCGTCAAGTA CATCCGCGCGCTGCA

SACHV3 -----

SACH CAAGAAGATGAGTAA GGTGAGACACTGCG CTCGGCGTCAAGTA CATCCGCGCGCTGCA

FIG.4J

1081

1140

SACHV1 GCAGCTGCTGGACGA GCATGACGGCGTGAG CGCCGCCCTTCCAGGC AGCGTCCGTGTCGCC
SACHV2 GCAGCTGCTGGACGA GCATGACGGCGTGAG CGCCGCCCTTCCAGGC AGCGTCCGTGTCGCC
SACHV3 -----TGCTGGACGA GCATGACGGCGTGAG CGCCGCCCTTCCAGGC AGCGTCCGTGTCGCC
SACH GCAGCTGCTGGACGA GCATGACGGCGTGAG CGCCGCCCTTCCAGGC AGCGTCCGTGTCGCC

1141

1200

SACHV1 CACCATCTCCCCCAA CTACTCCAACGACTT GAACCTCATGGCCGG CTCGCCGGTCTCATC
SACHV2 CACCATCTCCCCCAA CTACTCCAACGACTT GAACCTCATGGCCGG CTCGCCGGTCTCATC
SACHV3 CACCATCTCCCCCAA CTACTCCAACGACTT GAACCTCATGGCCGG CTCGCCGGTCTCATC
SACH CACCATCTCCCCCAA CTACTCCAACGACTT GAACCTCATGGCCGG CTCGCCGGTCTCATC

FIG.4K

1201

1260

SACHV1 CTACTCGTCGGACGA GGGCTCTTACGACCC GCTCAGCCCCGAGGA GCAGGAGCTTCTCGA
SACHV2 CTACTCGTCGGACGA GGGCTCTTACGACCC GCTCAGCCCCGAGGA GCAGGAGCTTCTCGA
SACHV3 CTACTCGTCGGACGA GGGCTCTTACGACCC GCTCAGCCCCGAGGA GCAGGAGCTTCTCGA
SACH CTACTCGTCGGACGA GGGCTCTTACGACCC GCTCAGCCCCGAGGA GCAGGAGCTTCTCGA

1261

1320

SACHV1 CTTCACCAACTGGTT CTGAGGGGCTCGGCC TGGTCAGGCCCTGGT GCCAATGGACTTTGG
SACHV2 CTTCACCAACTGGTT CTGAGGGGCTCGGCC TGGTCAGGCCCTGGT GCCAATGGACTTTGG
SACHV3 CTTCACCAACTGGTT CTGAGGGGCTCGGCC TGGTCAGGCCCTGGT GCCAATGGACTTTGG
SACH CTTCACCAACTGGTT CTGAGGGGCTCGGCC TGGTCAGGCCCTGGT GCCAATGGACTTTGG

FIG.4L

1321

1380

SACHV1 AAGCAGGGTGATCGC ACAACCTGCATCTTT AGTGCTTTCTTGTCA GTGGCGTTGGAGGG
SACHV2 AAGCAGGGTGATCGC ACAACCTGCATCTTT AGTGCTTTCTTGTCA GTGGCGTTGGAGGG
SACHV3 AAGCAGGGTGATCGC ACAACCTGCATCTTT AGTGCTTTCTTGTCA GTGGCGTTGGAGGG
SACH AAGCAGGGTGATCGC ACAACCTGCATCTTT AGTGCTTTCTTGTCA GTGGCGTTGGAGGG

1381

1440

SACHV1 GGAGAAAAGGAAAAG AAAA AAAAAGAAGAA GAAGAAGAAAAGAGA AGAAGAAA AAAACGA
SACHV2 GGAGAAAAGGAAAAG AAAA AAAAAGAAGAA GAAGAAGAAAAGAGA AGAAGAAA AAAACGA
SACHV3 GGAGAAAAGGAAAAG AAAA AAAAAGAAGAA GAAGAAGAAAAGAGA AGAAGAAA AAAACGA
SACH GGAGAAAAGGAAAAG AAAA AAAAAGAAGAA GAAGAAGAAAAGAGA AGAAGAAA AAAACGA

FIG.4M

1441

1500

SACHV1 AAACAGTCAACCAAC CCCATCGCCAACTAA GCGAGGCATGCCCTGA GAGACATGGCTTTCA
SACHV2 AAACAGTCAACCAAC CCCATCGCCAACTAA GCGAGGCATGCCCTGA GAGACATGGCTTTCA
SACHV3 AAACAGTCAACCAAC CCCATCGCCAACTAA GCGAGGCATGCCCTGA GAGACATGGCTTTCA
SACH AAACAGTCAACCAAC CCCATCGCCAACTAA GCGAGGCATGCCCTGA GAGACATGGCTTTCA

1501

1560

SACHV1 GAAAAACGGGAAGCGC TCAGAACAGTATCTT TGCACCTCCAATCATT CACGGAGATATGAAG
SACHV2 GAAAAACGGGAAGCGC TCAGAACAGTATCTT TGCACCTCCAATCATT CACGGAGATATGAAG
SACHV3 GAAAAACGGGAAGCGC TCAGAACAGTATCTT TGCACCTCCAATCATT CACGGAGATATGAAG
SACH GAAAAACGGGAAGCGC TCAGAACAGTATCTT TGCACCTCCAATCATT CACGGAGATATGAAG

FIG.4N

1561

1620

SACHV1 AGCAACTGGGACCTG AGTCAATGCCAATA TGCAGCTTGTGCA AAAGCAGTGGGCTCC
SACHV2 AGCAACTGGGACCTG AGTCAATGCCAATA TGCAGCTTGTGCA AAAGCAGTGGGCTCC
SACHV3 AGCAACTGGGACCTG AGTCAATGCCAATA TGCAGCTTGTGCA AAAGCAGTGGGCTCC
SACH AGCAACTGGGACCTG AGTCAATGCCAATA TGCAGCTTGTGCA AAAGCAGTGGGCTCC

1621

1680

SACHV1 TGGCAGAAGGAGCA GCACACGGCTTATAG TAACTCCCATCACC CTAAACAGCACAGCT
SACHV2 TGGCAGAAGGAGCA GCACACGGCTTATAG TAACTCCCATCACC CTAAACAGCACAGCT
SACHV3 TGGCAGAAGGAGCA GCACACGGCTTATAG TAACTCCCATCACC CTAAACAGCACAGCT
SACH TGGCAGAAGGAGCA GCACACGGCTTATAG TAACTCCCATCACC CTAAACAGCACAGCT

FIG.40

1681

1740

SACHV1 GAAAGTTCCTGCTCG GGTCCCTTCACCTCC CCGCCCTTCTTAGA GTGCAGTCTTAGCC
SACHV2 GAAAGTTCCTGCTCG GGTCCCTTCACCTCC CCGCCCTTCTTAGA GTGCAGTCTTAGCC
SACHV3 GAAAGTTCCTGCTCG GGTCCCTTCACCTCC CCGCCCTTCTTAGA GTGCAGTCTTAGCC
SACH GAAAGTTCCTGCTCG GGTCCCTTCACCTCC CCGCCCTTCTTAGA GTGCAGTCTTAGCC

1741

1800

SACHV1 CTCTAGAAACGAGTT GGTGCTTTCGCTTC AGTAGCCCCCACCAC AATAAGCTGTAGACA
SACHV2 CTCTAGAAACGAGTT GGTGCTTTCGCTTC AGTAGCCCCCACCAC AATAAGCTGTAGACA
SACHV3 CTCTAGAAACGAGTT GGTGCTTTCGCTTC AGTAGCCCCCACCAC AATAAGCTGTAGACA
SACH CTCTAGAAACGAGTT GGTGCTTTCGCTTC AGTAGCCCCCACCAC AATAAGCTGTAGACA

FIG.4P

1801

1860

SACHV1 TTGGTTTACAGTGAA ACTATGCTATTCTCA GCCCTTTGAAACTCT GCTTCTCCTCCAGGG
SACHV2 TTGGTTTACAGTGAA ACTATGCTATTCTCA GCCCTTTGAAACTCT GCTTCTCCTCCAGGG
SACHV3 TTGGTTTACAGTGAA ACTATGCTATTCTCA GCCCTTTGAAACTCT GCTTCTCCTCCAGGG
SACH TTGGTTTACAGTGAA ACTATGCTATTCTCA GCCCTTTGAAACTCT GCTTCTCCTCCAGGG

1861

1920

SACHV1 CCCGATTCCCAAAACC CCATGGCTTCCCTCA CACTGTCTTTTCTAC CATTTCATTATAGA
SACHV2 CCCGATTCCCAAAACC CCATGGCTTCCCTCA CACTGTCTTTTCTAC CATTTCATTATAGA
SACHV3 CCCGATTCCCAAAACC CCATGGCTTCCCTCA CACTGTCTTTTCTAC CATTTCATTATAGA
SACH CCCGATTCCCAAAACC CCATGGCTTCCCTCA CACTGTCTTTTCTAC CATTTCATTATAGA

FIG.4Q

1921

1980

SACHV1 ATGCTTCCAATCTTT TGTGAATTTTATTT ATAAAAAATCTATTT GTATCTATCCTAACC
SACHV2 ATGCTTCCAATCTTT TGTGAATTTTATTT ATAAAAAATCTATTT GTATCTATCCTAACC
SACHV3 ATGCTTCCAATCTTT TGTGAATTTTATTT ATAAAAAATCTATTT GTATCTATCCTAACC
SACH ATGCTTCCAATCTTT TGTGAATTTTATTT ATAAAAAATCTATTT GTATCTATCCTAACC

1981

2040

SACHV1 AGTTCGGGGATATAT TAAGATATTTTGTG CATAGAGAGAAGA GAGAGAAAAATTTAT
SACHV2 AGTTCGGGGATATAT TAAGATATTTTGTG CATAGAGAGAAGA GAGAGAAAAATTTAT
SACHV3 AGTTCGGGGATATAT TAAGATATTTTGTG CATAGAGAGAAGA GAGAGAAAAATTTAT
SACH AGTTCGGGGATATAT TAAGATATTTTGTG CATAGAGAGAAGA GAGAGAAAAATTTAT

FIG.4R

2041

2100

SACHV1 AGAAGTTTGTACAA ATGGTTTAAATGTG TATATCTTGATACTT TAACATGTAATGCTA
SACHV2 AGAAGTTTGTACAA ATGGTTTAAATGTG TATATCTTGATACTT TAACATGTAATGCTA
SACHV3 AGAAGTTTGTACAA ATGGTTTAAATGTG TATATCTTGATACTT TAACATGTAATGCTA
SACH AGAAGTTTGTACAA ATGGTTTAAATGTG TATATCTTGATACTT TAACATGTAATGCTA

2101

2160

SACHV1 TTACCTCTGCATATT TTAGATGTGTAGTTC ACCTTACAACCTGCAA TTTTCCCTATGTGGT
SACHV2 TTACCTCTGCATATT TTAGATGTGTAGTTC ACCTTACAACCTGCAA TTTTCCCTATGTGGT
SACHV3 TTACCTCTGCATATT TTAGATGTGTAGTTC ACCTTACAACCTGCAA TTTTCCCTATGTGGT
SACH TTACCTCTGCATATT TTAGATGTGTAGTTC ACCTTACAACCTGCAA TTTTCCCTATGTGGT

FIG.4S

2161

2220

SACHV1 TTTGTAAGAAGACTCT CCTCATAGGTGAGAT CAAGAGGCCACCAGT TGTACTTCAGCACCA

SACHV2 TTTGTAAGAAGACTCT CCTCATAGGTGAGAT CAAGAGGCCACCAGT TGTACTTCAGCACCA

SACHV3 TTTGTAAGAAGACTCT CCTCATAGGTGAGAT CAAGAGGCCACCAGT TGTACTTCAGCACCA

SACH TTTGTAAGAAGACTCT CCTCATAGGTGAGAT CAAGAGGCCACCAGT TGTACTTCAGCACCA

2221

2280

SACHV1 ATGTGTCTTACTTTA TAGAATGTTGTTAA TGTATTAAATGATGTT ATTAAATACTGTCA

SACHV2 ATGTGTCTTACTTTA TAGAATGTTGTTAA TGTATTAAATGATGTT ATTAAATACTGTCA

SACHV3 ATGTGTCTTACTTTA TAGAATGTTGTTAA TGTATTAAATGATGTT ATTAAATACTGTCA

SACH ATGTGTCTTACTTTA TAGAATGTTGTTAA TGTATTAAATGATGTT ATTAAATACTGTCA

FIG.4T

2281

2340

SACHV1 AGAAGACAAAGTTT ATGCAGCTACTGTCC AAACCTCAAAGTGCCA GCCAGTTGGTTTGA
SACHV2 AGAAGACAAAGTTT ATGCAGCTACTGTCC AAACCTCAAAGTGCCA GCCAGTTGGTTTGA
SACHV3 AGAAGACAAAGTTT ATGCAGCTACTGTCC AAACCTCAAAGTGCCA GCCAGTTGGTTTGA
SACH AGAAGACAAAGTTT ATGCAGCTACTGTCC AAACCTCAAAGTGCCA GCCAGTTGGTTTGA

2341

2400

SACHV1 TAGGTTGCCCTTTGG AGATTCTATTACTG CCTTTTCTTCTTAC TGTTTATTACAAC
SACHV2 TAGGTTGCCCTTTGG AGATTCTATTACTG CCTTTTCTTCTTAC TGTTTATTACAAC
SACHV3 TAGGTTGCCCTTTGG AGATTCTATTACTG CCTTTTCTTCTTAC TGTTTATTACAAC
SACH TAGGTTGCCCTTTGG AGATTCTATTACTG CCTTTTCTTCTTAC TGTTTATTACAAC

FIG.4U

2401

2460

SACHV1 TTACAAAAATATGTA TAACCCCTGTTTATA CAAACTAGTTTCGTA ATAAAACTTTTTCCT
SACHV2 TTACAAAAATATGTA TAACCCCTGTTTATA CAAACTAGTTTCGTA ATAAAACTTTTTCCT
SACHV3 TTACAAAAATATGTA TAACCCCTGTTTATA CAAACTAGTTTCGTA ATAAAACTTTTTCCT
SACH TTACAAAAATATGTA TAACCCCTGTTTATA CAAACTAGTTTCGTA ATAAAACTTTTTCCT

2461

SACHV1 TTTTTTTAAATG 1960
SACHV2 TTTTTTTAAATG 1994
SACHV3 TTTTTTTAAATG 2304
SACH TTTTTTTAAATG 2472

60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

SACHV MESSAKMESGGAGQÖ PÖPÖPÖQÖPTLPAC FTATTAATAAAAAAAAAA AAQSAQÖQÖQÖQÖQÖQÖ

120

-----MRCKRLNE SFGYSLPQQPAAV

SACHV QQAPQLRPAADGÖPS GGGHKSAPKÖVKRÖR SSSPEIMRCKRRLNF SGFGYSLPQQÖPAAY

FIG.5B

121

180

SACHV1 ARRNERERNRVKLVN LGFATLREHVPNGAA NKKMSKVETLRSAVE YIRALQQLDEHDAV
 SACHV2 -----MSKVETLRSAVE YIRALQQLDEHDAV
 SACHV3 -----LLDEHDAV
 SACHV ARRNERERNRVKLVN LGFATLREHVPNGAA NKKMSKVETLRSAVE YIRALQQLDEHDAV

181

SACHV1 SAAFOAGVLSPTISP NYSNDLNSMAGSPVS SYSSDEGSYDPLSPE EQELLDFTNWF 140
 SACHV2 SAAFOAGVLSPTISP NYSNDLNSMAGSPVS SYSSDEGSYDPLSPE EQELLDFTNWF 83
 SACHV3 SAAFOAGVLSPTISP NYSNDLNSMAGSPVS SYSSDEGSYDPLSPE EQELLDFTNWF 180
 SACHV SAAFOAGVLSPTISP NYSNDLNSMAGSPVS SYSSDEGSYDPLSPE EQELLDFTNWF 236